AD_____

Award Number: DAMD17-03-1-0124

TITLE: 3D Structure Determination of a Human Prostate-Specific Homeoprotein, NKX3.1: The Mechanism of Autoregulation Through its Terminal Domains

PRINCIPAL INVESTIGATOR: York Tomita, Ph.D.

CONTRACTING ORGANIZATION: Georgetown University Medical Center
Washington, DC 20007

REPORT DATE: February 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 2004	Annual (1 Feb 2003 - 31 Jan 2004)

4. TITLE AND SUBTITLE
3D Structure Determination of a Human Prostate-Specific Homeoprotein, NKX3.1: The Mechanism of Autoregulation Through its Terminal Domains

5. FUNDING NUMBERS
DAMD17-03-1-0124

6. AUTHOR(S)
York Tomita, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Georgetown University Medical Center
Washington, DC 20007

E-Mail: yat@gunet.georgetown.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Original contains color plates: All DTIC reproductions will be in black and white.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)
NKX3.1, a member of the NK class of homeodomain (HD) proteins, is expressed primarily in the adult prostate and has growth suppression and differentiating effects in prostate epithelial cells. NKX3.1 consists of three domains having the HD in the middle. The NKX3.1 HD is known to binds to DNA for its transcriptional activity and other transcriptional factors such as serum response factor (SRF). Our goal is to determine 3D structure of NKX3.1 to elucidate the mechanism of its functions and regulations.

During the first year, full-length as well as various truncation constructs of NKX3.1, including N-terminal deletions, C-terminal deletions, and HD only, were made and purified in a sufficient quantity for the structural characterization by NMR and circular dichroism (CD). The interaction between NKX3.1 HD and DNA was confirmed by EMSA and NMR, and the binding of SRF to NKX3.1 HD was confirmed by NMR. Using the 13C/15N double labeled NKx3.1 HD in complex with DNA, a suite of 3D NMR experiments were recoded for the backbone and sidechain resonance assignments. The spectral analysis is in progress and soon we will be able to calculate 3D NMR structures of the complex. We will continue our structural study with full-length NKX3.1.

14. SUBJECT TERMS
NKX3.1, structure, NMR

15. NUMBER OF PAGES
12

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

I. Introduction

The homeodomain of NKX3.1 shares high sequence homologies with a large family of homeobox proteins that hold conserved 3D structures. The closest known 3D structure of the NKX3.1 homeodomain is the homeodomain of the drosophila VND/NK-2 protein [1, 2] with amino acid sequence identity over 60%. In additional to the homeodomain, the NK-2 class of proteins has two highly conserved regions, N-terminal TN (tin) domain (NK decapeptide), and the NK2-specific domain (NK2-SD or the NK2 domain). So far 3D structure of the entire NK-2 is not yet known due to low expression and limited solubility of the proteins. Although NKX3.1 and NK-2 share a high degree of sequence homology within the homeodomain, the identity between NKX3.1 and NK-2 drops down to less than 10% for the regions out side of the homeodomain. NKX3.1 does not have the TN domain or NK2-SD. Instead, NKX3.1 has shorter and more hydrophilic C-terminus than NK-2. Therefore, NKX3.1 may be more amenable to structure characterization of the full-length protein. In fact, we have succeeded in obtaining large amounts of electrophoretically pure full-length and deletion constructs of NKX3.1 This has allowed us to generate preliminary magnetic resonance spectra to determine the solution structure of this full-length homeodomain protein. This is an extremely important finding for prostate cancer and a very significant development since this would be the first homeodomain protein for which a complete solution structure would be determined. In addition to binding DNA via the homeodomain, NKX3.1 also binds to other transcription factors such as serum response factor (SRF). For these reasons it is clear that the actions of NKX3.1 are complex and its interactions with other transcriptional complexes such as those that are initiated by AR and by the β-catenin/TCF complex may underlie its biological activity and tumor suppressor role.

II. Body

Specific Aim Proposed for This Time Period

Aim 1 3D Structure Determination of Full-Length NKX3.1 by Solution NMR (18 months)

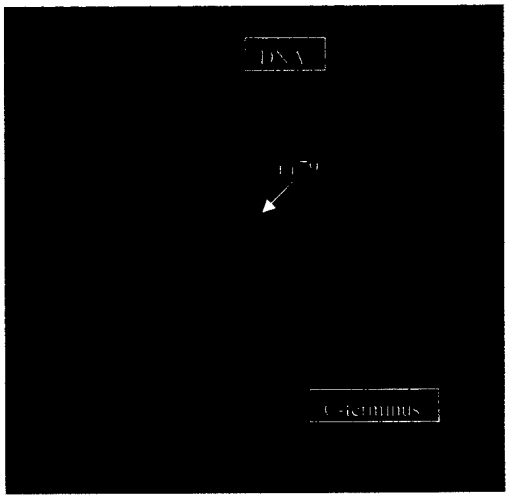
NKX3.1 has shorter and more hydrophilic C-terminus than NKX2.5 or NK-2 that it may be more amenable to structure characterization for the full-length protein. The determination of full-length NKX3.1 structure will unambiguously verify our hypothesis that the interaction between the two termini prevents the NKX3.1DNA binding activities. We will use the following experimental approach.

a. Engineer and express full-length NKX3.1 in soluble form for the structural study (6 months)
b. High-resolution NMR experiments on the recombinant protein to determine the 3D structure of NKX3.1 (8 months)
c. 3D structure calculation with X-PLOR and structural analysis (4 months)

Model Structure of NKX3.1

A model structure of NKX3.1 bound to DNA was built using Insight II / modeler interface (Accelrys. Inc.) (Figure 1) based on the NMR 3D structure of NK-2 in complex with DNA (pdb code: 1NK3) [1]. In this model the N- and C-termini are shown in random coils since no suitable known 3D structures to build the models have been found. Therefore, experimental elucidation of the 3D structure of the full-length NKX3.1 by NMR is desirable to understand the interactions of the homeodomain and the two terminal domains. The model structure at the NKX3.1 homeodomain is expected to be reasonable

Figure 1 A Model Structure of NKX3.1 Bound to DNA



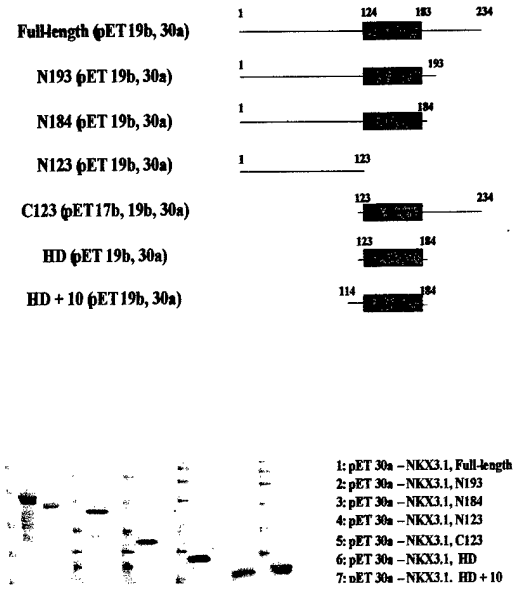
The NK-2 protein (blue ribbon) bound to its DNA (coral wireframe) and a model of NKX3.1 (green). The T179 residue of NKX3.1 is shown (red wireframe) and corresponding NK-2 residue (also a T - blue wireframe) are shown.

because of its high sequence identity to NK-2, but still it fails to explain a number of observations. For example, 1) why NK-2 and NKX3.1 have different specificities in their DNA sequence recognitions, and 2) why the residue T179 is critical for DNA recognition, but not for SRF binding by NKX3.1, while the corresponding residue in NK-2 is solvent exposed and it is not participating in the DNA interaction.

Expression of Recombinant NKX3.1 (aim1-a)

The recombinant full-length NKX3.1 was overexpressed in *E. coli* as a polyhistidyl-tagged fusion protein. The fusion-protein was purified from cytosolic fraction (soluble fraction) using Ni affinity column chromatography (Ni-NTA resin), followed by SP ion-exchange chromatography. The full-length NKX3.1 was soluble and remained soluble all the time. The expression level of the protein was also much more favorable than NK-2 (total yield, ~1mg of purified protein form 1L of cell culture).

Figure 2 Expressions of Various Truncation Mutants of NKX3.1



In addition to the full-length protein, several truncated versions of the NKX3.1 proteins were also constructed and purified in the similar method as the full-length protein (figure 2). Unlike its homologue NK-2, all forms of NKX3.1 were expressed well and stayed as soluble forms throughout the purification and thereafter (total typical yield, ~3mg of purified protein form 1L of cell culture).

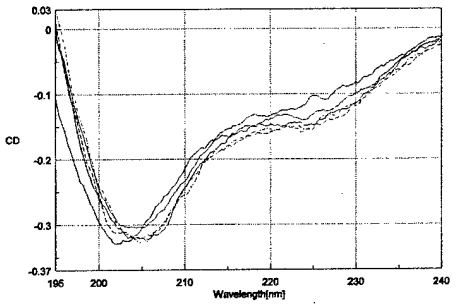
The ^{15}N labeled forms of the proteins were also produced for full-length NKX3.1, NKX3.1(Δ184-234), and NKX3.1(Δ1-122) using ^{15}N ammonium chloride as a sole source of nitrogen in M9 media. The final ^{15}N labeled protein yield after the purification was approximately three mg per liter of cell culture, which ensures that enough protein in isotope labeled forms can be produced for the NMR structural studies. The polyhistidyl

tag was removed by thrombin digestion and gel filtration column and/or ion-exchange column were applied. Similarly, $^{13}C/^{15}N$ double labeled NKX3.1 HD+10 (114-184) was produced for a full structure characterization by NMR.

Secondary Structure Analysis of NKX by Circular Dichroism

The secondary structural content and thermal stability of NKX3.1 was tested with circular dichroism (CD) spectroscopy. Molar ellipticity at 222 nm showed that the full-length NKX3.1 and NKX3.1(Δ184-234) proteins have α-helical structure, which is expected from the homeobox proteins (figure 3). The thermal unfolding of the NKX3.1(Δ184-234) protein was monitored by changes in the molar ellipticity at 222 nm. The results showed that the protein was stable up to 40°C and above this temperature, unfolding of the protein occurred (figure 3). The unfolding event was reversible when the temperature was returned to below 40°C.

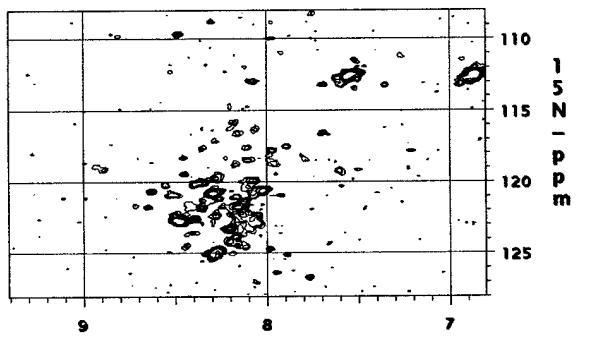
Figure 3 Thermal Stability of NKX3.1(Δ184-234) by CD



NMR spectra of NKX3.1 (aim1-b)

Using ^{15}N-labeled proteins of full-length and truncated NKX3.1, ^{15}N-heteronuclear single-quantum correlation (HSQC) NMR spectra were recorded (full-length NKX3.1 spectrum shown in Figure 4). ^{15}N HSQC spectra are referred to as "fingerprints" of proteins because they reflect unique 3D structures of the proteins. The HSQC spectra of NKX3.1 showed

Fig 4 N-HSQC Spectrum of NKX3.1 at 25 C



dispersed peaks and it appeared to be folded protein with a high α-helical content and some flexible segments as consistent with the CD results. The 3D protein structures seemed to be stable at wide range of conditions including pH, salt concentration and temperature (up to 40°C). The results from the HSQC spectra are encouraging for pursuing the NMR structure elucidating of full-length NKX3.1.

The NKX3.1 Proteins Binding to the Consensus DNA oligonucleotide

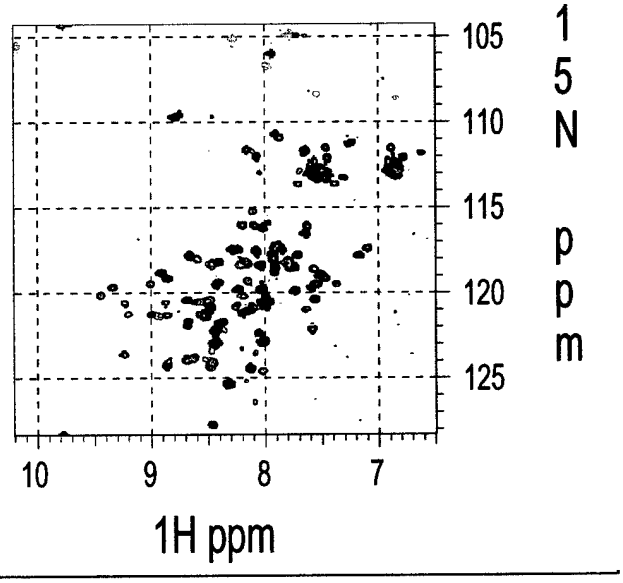
Previously we identified the TAAGTA consensus binding site for NKX3.1 and, demonstrated that NKX3.1 preferentially binds to the TAAGTA sequence rather than the consensus binding site for NKX2.1 (CAAGTG) or MSX1 (TAATTG) [3]. We designed six short DNA oligonucleotides, containing the TAAGTA site in the middle (14 and 18 mers).

```
  -5-4-3-2-1 1 2 3 4 5 6 7 8 9
5'-G G G * * T A A G T A C C C-3' (14-mers)

  -7-6-5-4-3-2-1 1 2 3 4 5 6 7 8 9 10 11
5'-* * G G G * * T A A G T A C C C  *  *-3' (18-mers)
```

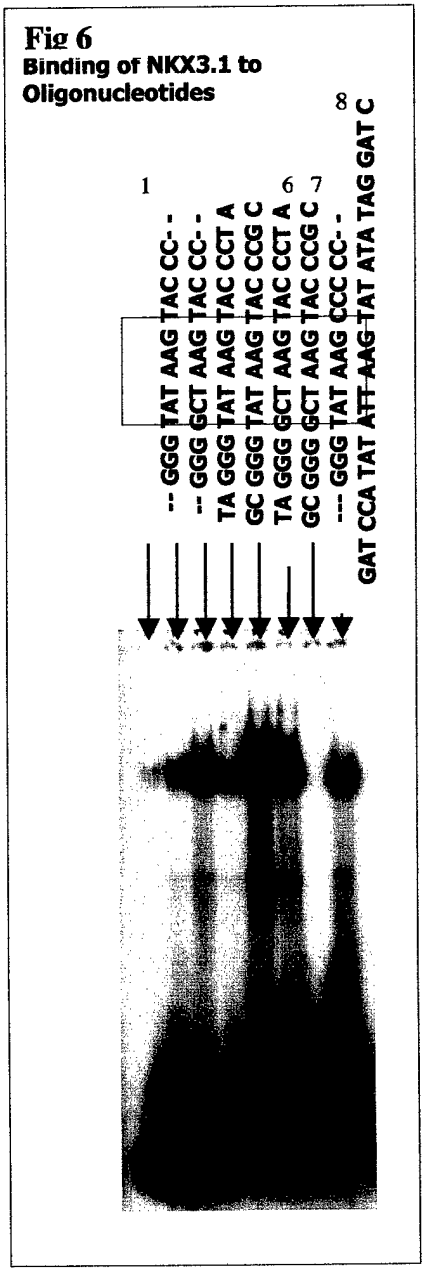
where ** is either GC or TA. A seventh oligo contained the core TAAGCC that has no binding to NKX3.1 (a negative control) and an eighth oligo is a positive control [3] (Figure 6).

Fig 5 ^{15}N HSQC spectrum of NKX3.1 HD+10 (114-184) bound to DNA at 30 deg C



Comparing the 3D structure of the NK-2/DNA complex with the sequence of NKX3.1 (Figure 1), the nuleotides directly in contact with the NKX3.1 homeodomain are expected to be from positions 1 to 5

(TAAGT). In addition, the N-terminal residues of NKX3.1 may interact with DNA at positions from -2 to 1 (**T). All six duplex oligonucleotides with the core consensus sequence bound to recombinant NKX3.1 as shown by electromobility shift assay, EMSA (Figure 6). We also recorded ^{15}N-HSQC spectra with NKX3.1 HD+10 (114-184) and one of the 14-mer DNA constructs (#2) (Figure 5) and one of the 18-mer DNA constructs (#6) (not shown). Both oligos made stable complexes with NKX3.1 and caused induced chemical shifts in many resonance positions including side chain amines (Figure 5). In figure 5, the black spectrum is free NKX3.1 HD+10, and the red spectrum is NKX3.1 H+10 in complex with DNA.



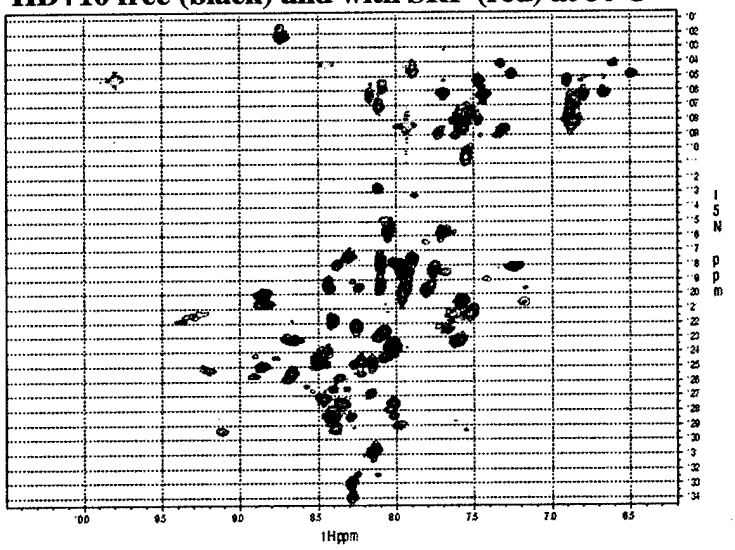
Mapping the interface between NKX3.1 and SRF

NKX3.1 has been shown to interact with SRF and enhances its transcription activity. The interaction domains are mapped to the homeodomain in NKX3.1 and the MADS box in SRF [4], and the protein-protein interaction is independent of DNA. We purified the MADS box of SRF and tested its direct binding with NKX3.1 using ^{15}N HSQC spectroscopy. When none-labeled SRF was titrated into ^{15}N labeled NKX3.1 homedomain (NKX3.1 HD+10, residues 114-184), the spectra showed many induced chemical shifts, confirming NKX3.1 HD and SRF MADS box indeed interacted directly (Figure 7). Once the resonance assignments of the backbone amides become available during the second year of this project, the binding site of SRF on the 3D structural surface of NKX3.1 will be reviled.

The interaction between NKX3.1 and SRF seems to be transient as the ternary complex (SRF/NKX3.1/DNA) was not observed in DNA Binding Mobility Shift Assay [4]. NMR is an ideal tool to study weak

inter-molecular interactions in solution and we will determine the binding constant between NKX3.1 and SRF by NMR. Similarly ternary interactions among NKX3.1, SRF and DNA will be analyzed. Finally, model complex structures of the ternary complex will be built using molecular docking with NMR derived restraints [5].

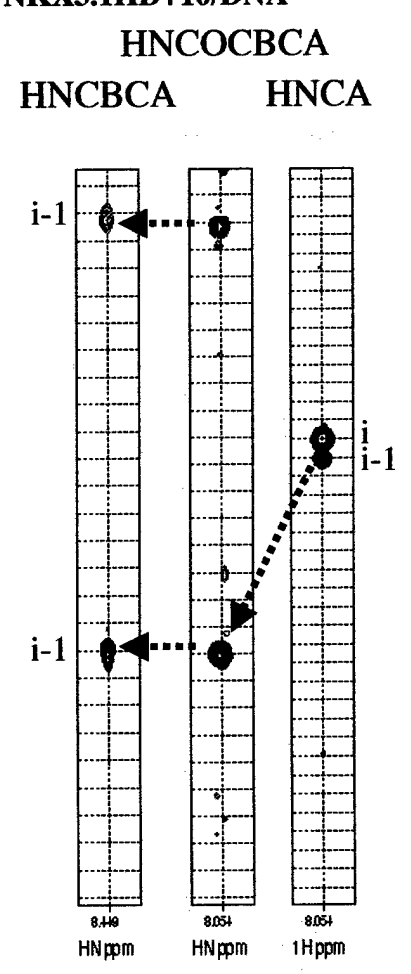
Figure 7 Overlaid ^{15}N HSQC spectra of NKX3.1 HD+10 free (black) and with SRF (red) at 30 C



Progress of 3D Structure Determination of NKX3.1

NKX3.1 consists of the three domains, and the homeodomain (HD) can fold independently from the rest of the protein. To simplify spectral analysis, we decided to characterize the 3D structure of the HD in complex with DNA first. Using ^{13}C/^{15}N double labeled NKX3.1HD+10 (residues 114-184), a standard suite of triple resonance experiments [6] (3D HNCA, HNCBCA, HNCOCBCA, HNCO, HCACO, 3D C(CO)NH) have been recorded for the resonance assignments of backbone and side-chain atoms for NKX3.1 HD+10 in complex with DNA. The protein complex was stable throughout the experiments and gave well dispersed peaks with good intensities in the 3D spectra (figure 8). We are currently analyzing the spectra to complete the backbone assignments and soon move on to the side-chain assignments.

Figure 8 Strips of 3D NMR spectra of NKX3.1HD+10/DNA



III Key Research Accomplishment

Progress on our project has been excellent and additional findings on details of the DNA binding and SRF interaction with NKX3.1 will be forthcoming shortly. During the first period, we have accomplished:

1. Constructed and purified full-length and various truncation mutants of NKX3.1.
2. ^{15}N labeled and $^{13}C/^{15}N$ labeled ^{15}N proteins haven been purified for the NMR studies.
3. ^{15}N HSQC and CD spectra were obtained for some of the above proteins to asses the secondary structural contents and thermal stabilities.
4. Complex formations were confirmed with various constructs of DNA and NKX3.1 by EMSA and NMR.
5. Direct interaction between NKX3.1 and SRF was confirmed by NMR
6. Triple resonance 3D NMR experiments have been recoded with the NKX3.1 homedodoamin/DNA complex, and the resonance assignments are in progress.

IV. Reportable Outcome

None

V Conclusions

Unlike its homologue NK-2, we were able to express and purify a large amount of NKX3.1 as a stable and soluble form. Therefore it will open an excellent opportunity to elucidate the 3D structure of the full-length protein for the first time in this class of proteins. During the first year, full-length as well as various truncation constructs of NKX3.1, including N-terminal deletions, C-terminal deletions, and HD only, were made and purified for the structural characterization by NMR and circular dichroism (CD). Both methods showed that the homeodomain is well folded with high α-helical contents and there are some disordered regions in other

domains. Overall the NMR spectra were reasonably well dispersed and it seems to be amenable to NMR structural studies.

The interactions between NKX3.1 HD and DNA were confirmed by EMSA and NMR, and showed that as short as 14-mer double helical DNA with NKX3.1 consensus site was able to bind to NKX3.1. The binding of serum response factor (SRF) to NKX3.1 HD was confirmed by NMR. It shows that NKX3.1 homeodomain was able to bind directly to the MADS box of SRF. DNA induced large conformational changes throughout the homeodomain, while SRF interaction was contained in a small area of NKX homeodomain.

Using the $^{13}C/^{15}N$ double labeled NKX3.1 HD in complex with DNA, a suite of 3D NMR experiments was recoded for the backbone and sidechain resonance assignments. Most of the peaks are well resolved in the3D spectra, and the spectral analysis is in progress smoothly. We will be able to calculate 3D NMR structures of the complex soon. Once the homeodomain structure is characterized, we will continue our structural study with full-length NKX3.1.

"So what?": We have made a significant progress in solving the 3D structure of the NKX3.1. The 3D structure will reveal detailed molecular interactions between NKX3.1 and DNA, regulation of NKX3.1 function, protein-protein interactions with other factors such as SRF.

VI References

1. Gruschus, J.M., et al., *Interactions of the vnd/NK-2 homeodomain with DNA by nuclear magnetic resonance spectroscopy: basis of binding specificity.* Biochemistry, 1997. **36**(18): p. 5372-80.
2. Gruschus, J.M., et al., *The three-dimensional structure of the vnd/NK-2 homeodomain-DNA complex by NMR spectroscopy.* J Mol Biol, 1999. **289**(3): p. 529-45.
3. Steadman, D.J., D. Giuffrida, and E.P. Gelmann, *DNA-binding sequence of the human prostate-specific homeodomain protein NKX3.1.* Nucleic Acids Res, 2000. **28**(12): p. 2389-95.
4. Carson, J.A., et al., *The smooth muscle gamma-actin gene promoter is a molecular target for the mouse bagpipe homologue, mNkx3-1, and serum response factor.* J Biol Chem, 2000. **275**(50): p. 39061-72.
5. Clore, G.M. and C.D. Schwieters, *Docking of protein-protein complexes on the basis of highly ambiguous intermolecular distance restraints derived from 1H/15N chemical shift mapping and backbone 15N-1H residual dipolar couplings using conjoined rigid body/torsion angle dynamics.* J Am Chem Soc, 2003. **125**(10): p. 2902-12.
6. Sattler, M., J. Schleucher, and C. Griesinger, *Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulse field gradient.* Progress in NMR spectroscopy, 1999. **34**: p. 93-158.